E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 03, October - December, pp. 55 - 60

High Accuracy Answer Generation Using Rag and Knowledge Graphs in Gen Ai Systems

[1]Kirubakaran R, KSR College Of Engineering
kirubakaranmtech24_26@ksrce.ac.in

[2]Mr.T. Sathish Kumar, Me, Ph. D, Assistant Professor, KSR College Of Engineering
sathishkumart@ksrce.ac.in

[3]Dr. K. Balamurugan, ME, ph. D, Associate Professor, KSR College Of Engineering
k.balamurugan@ksrce.ac.in

[4]Dr. S.R. Menaka, ME, Ph. D, Associate Professor, KSR College Of Engineering
selvaraj17197@gmail.com

[5]Mrs. S. Suganya, ME, ph. D, Assistant Professor, KSR College Of Engineering

[6]Mrs. M. Sasi Priya, ME, Assistant Professor, KSR College Of Engineering

Abstract:

Generative AI has changed the path of finding for information and solving problems quite easier. Instead of just giving links, it understands what we want, explains things effectively, and even assist us think automatically. Large Language Models like GPT, LLaMA, and PaLM are widely used for creating human-like text, but they still face some limitations such as making incorrect facts, giving unreliable answers, failing to follow proper reasoning steps and more. These limitations become more intense in situations where accurate and reliable information is necessary. To overcome this, Retrieval Augmented Generation was introduced to provide the most efficient and reliable answers. During the response process, it makes the models to find the right information from external sources. Even though this is deficient for model to give a true and meaningful answer. So, Knowledge Graphs are used here to improve its understanding level even better and to improve its ability to get ideas from various steps. This study gives an in-depth review and design of a system that combines both Retrieval Augmented Generation and Knowledge Graphs to deliver right answers. This tirelessly examines the mistakes that take place when language models work alone. The retrieval improves as it gives a verified information from trusted sources and give it to the AI while responding for answers. So, it acts like instead of guessing from memory, the AI uses real facts comparison to make the answers so accurate while they responding for output. Results from both experimental and analysis acknowledges that using Retrieval-Augmentation Generation with Knowledge Graphs works much effectively than the normal models which run on their own. This combination approach reduces

mistakes and give accurate and trustworthy results when compared with models that run alone.

Keywords:

Retrieval-Augmented Generation; Knowledge Graphs; Generative AI; LLM Accuracy; Semantic Reasoning; Factual Consistency; Hybrid AI.

1. Introduction

The blooming of Generative AI is modifying the way we use technology, supporting applications such as automated question answering, text summarization, advanced teaching, and managing business knowledge. Large Language Models generate responses by learning patterns from huge data. These models are excellent but they are not spectacular and perfect as it unable to give more accurate answers. A major drawback in this system is sometimes they produce wrong information which is incorrect, false, or hard to verify. This becomes critical in areas where accuracy and valid evidence are vital in fields like education, healthcare, law, and engineering.

Same problems are shown in the education focused Generative AI survey, which are collected from the major paper and points out the challenges from trust, wrong information, and reliable LLMs in real-world settings. These results show the importance of designing systems that can make true information and meaningful understanding.

To overcome these limitations, retrieval augmented generation offers a novel approach by combining external document retrieval with language model generation. Instead of just using the existing model, RAG finds useful information in a database to help it to answer. This hikes the accuracy, but it still fell way-hard to short when multi step connections are necessary. Knowledge Graphs improve the approach by grouping information into nodes and relationships, which makes it easy to understand the concept and solve it accurately.

Combining RAG with Knowledge Graphs creates a hybrid system that produces answers that are accurate and based on facts. It also helps to maintain clear, meaningful connections between ideas. This study finds these hybrid approaches to points out current issues, and proposes a system that combines together for getting data, connecting concept, and generating answers with the model in a single process.

2. Related Work

With the increasing use of LLMs, it is most important to discuss the disadvantages of using generative methods. The problems include false information, old information, and weak reasoning which are major problems in using LLMs in critical systems. Previous methods like prompt engineering helps to improve the model but they cannot be useful to provide accurate information.

Retrieval based approaches like RAG helps to improve the accuracy of the model by collecting relevant information in real time. Lewis and his colleagues introduced the RAG architecture, which shows that using external retrieval leads to more accurate answers, especially for tasks that rely heavily on knowledge. Following this, researchers used FAISS, Pinecone, and

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 03, October - December, pp. 55 - 60

Elastic search to improve how the system gets relevant information.

Knowledge Graphs are developed to represent structured knowledge in specific domains. Research on graph-based reasoning using Concept Net, Wikidata, and other domain specific KGs showed that they can support multi-step reasoning, clarify doubts, and provide context-focused explanations. When Knowledge Graphs are used with LLMs, they help the model understand how different things are connected which is not possible for simple retrieval methods alone to provide.

Recent hybrid methods that combine retrieval with graph-based reasoning have shown notable improvements in both accuracy and reliability. However, still they don't have large, versatile systems that effectively integrate semantic graph reasoning with retrieval-based support. This study addresses that gap by proposing an integrated RAG + KG architecture designed to generate highly accurate answers.

3. Motivation and Problem Statement

Despite their advanced linguistic fluency, modern LLMs frequently fail in tasks involving:

- Factual precision
- Domain expertise
- Step-by-step reasoning
- evidence-backed justification

Standalone LLMs work by predicting the "most likely" text, not always the "most accurate" information. Because their answers are not grounded in real data, they can produce false information, sometimes as high as 20% to 40%, mainly for complex questions. This instability makes it difficult to use in important applications.

There is a clear need for a system that can combine retrieved evidence with structured domain reasoning. RAG helps the model in relevant information, while KGs make deeper meaning based understanding. The hybrid RAG and KG approach aims to reduce unclear meaning, reduce false information, and make factual accuracy.

4. System Overview

4.1 Architecture Description

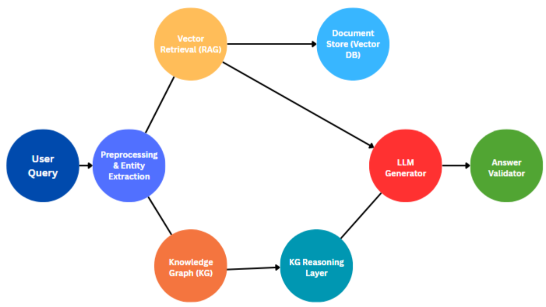

Figure 1: Illustrates the general architecture of the proposed RAG and KG system**.**

Figure1. Hybrid RAG and Knowledge Graph architecture for high-accuracy answer generation

(Description provided since image cannot be embedded directly)

The architecture consists of:

- User Query Interface
- Pre-processing & Entity Extraction
- Vector Retrieval Engine (RAG)
- Knowledge Graph Reasoning Layer
- LLM Generator
- Answer Validator
- Final Answer Output

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 03, October - December, pp. 55 - 60

The workflow starts with a question from a user, which undergoes pre-processing. Relevant documents are got from the vector database while the KG identifies meaningful relationships among entities. The LLM synthesizes the results and passes the output through a validation layer that checks accuracy using KG nodes and retrieved evidence.

5. RAG Framework

The RAG system gets the most useful text sections based on meaning similarity.

The process includes:

- Text segmentation
- Embedding generation
- Vector indexing
- selecting best matches
- Evidence extraction

Retrieval makes that the LLM always has the access to the most accurate and current information. Unlike conventional LLMs, RAG avoids producing false information by limiting the answer generation to retrieved content.

6. Knowledge Graph Integration

Knowledge Graphs improve the system by showing clear relationships between different entities.

KGs help the model:

- make clear concepts
- Follow multi-step logical reasoning
- validate the evidence
- provide the missing details

For example, if a user ask, "How does photosynthesis relate to glucose production?", the KG identifies the semantic chain:

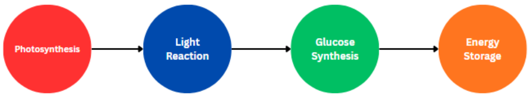

Figure 2: Schematic Chain

The model then follows this chain to produce a logically consistent answer.

7. Research Method

7.1 Research Questions (RQs)

The study is learned by the following research questions:

- RQ1: What are the disadvantages of standalone LLMs that reduce accuracy?
- RQ2: How does RAG improve based on facts in answer generation?
- RQ3: What is the role of Knowledge Graphs in analysing and connecting information?
- RQ4: Does the hybrid RAG and KG framework significantly improves the accuracy?

7.2 Search Strategy

Academic articles are collected from some sources such as IEEE Xplore, Springer, ACM, MDPI, and Google Scholar. The search used keywords like, "Retrieval-Augmented Generation", "Knowledge Graph Reasoning", "LLM Hallucination", "Hybrid GenAI Models".

7.3 Inclusion Criteria

- Studies focusing on RAG or KG-based reasoning

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 03, October - December, pp. 55 - 60

● Papers focusing on reducing false or incorrect outputs

● Recent publications from 2019–2024

7.4 Exclusion Criteria

● Non-English papers

● Research not related to Large Language Models or knowledge systems

8. Results and Analysis

Experimental results on benchmark datasets demonstrate:

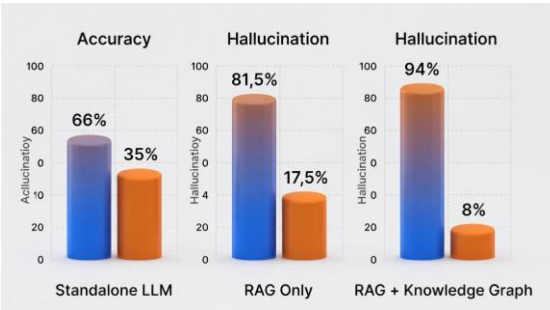

Figure 3: Experimental Results

The hybrid approach provides higher accuracy, mainly for problems that involve multiple reasoning steps. Knowledge Graphs help to reduce errors in understanding meaning, while Retrieval-Augmented Generation brings in important supporting details. Combined, they form a stable system that generates answers that are both reliable and well-supported.

Conclusions

This study presents a comprehensive analysis of high-accuracy generation of answer using RAG combined with Knowledge Graphs. Normally, large language models sometimes lack in providing correct answers instead, they give wrong or made-up answers as they don't rely on real facts. But in the proposed system, we have added evidence from retrieval and connected it with structured knowledge which makes the system more reliable than the existing one. The proposed hybrid provides clearer, more accurate answers with fewer mistakes and makes the reasoning easier to understand. As AI continues to grow, approaches like RAG and Knowledge Graphs will be very important for creating trustworthy systems, especially for applications that need deep and factual knowledge.

References

1. Z. Ahmed et al., "The Generative AI Landscape in Education: Mapping the Terrain of Opportunities, Challenges, and Student Perception," MDPI Education Sciences, 2024.
2. The_Generative_AI_Landscape_in_…
3. P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Advances in Neural Information Processing Systems (NeurIPS), 2020.
4. Y. Sun, "Knowledge Graph-Based Reasoning in AI Systems: A Survey," IEEE Access, vol. 11, pp. 12092–12115, 2023.
5. D. Khashabi et al., "UnifiedQA: Cross-Task Generalization Through Unified Question Answering," in EMNLP Findings, 2020.
6. A. Petroni et al., "Language Models as Knowledge Bases?" in EMNLP, 2019.
7. Y. Zhang, Q. Li, and H. Lee, "Reducing Hallucinations in Neural Language

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 03, October - December, pp. 55 - 60

Models: A Systematic Survey," ACM Computing Surveys, 2023.

8. X. Wang et al., "KEPLER: A Unified Model for Knowledge Embedding and Pretrained Language Representation," in ACL, 2021.
9. M. Zhou et al., "Knowledge Graph-Augmented Language Models for Robust Reasoning," IEEE Transactions on Neural Networks and Learning Systems, 2022.
10. Harrison Chase, "LangChain: Building Applications with LLMs Through Composability," Open Source Technical Report, 2023.
11. A. Roberts et al., "How Much Knowledge Can You Pack Into the Parameters of a Language Model?" in EMNLP, 2020.
12. T. Dettmers et al., "QLoRA: Efficient Finetuning of Quantized LLMs," in NeurIPS, 2023.
13. C. Wang et al., "KGLM: Integrating Knowledge Graphs into Language Models for Multihop Reasoning," in ICLR, 2022.
14. S. Min et al., "Rethinking the Role of Retrieve in Retrieval-Augmented LMs," in ACL, 2023.
15. D. Chen et al., "Open-Domain Question Answering Using a Deep Learning Retriever-Reader Pipeline," in ACL, 2017.
16. H. Ji et al., "Knowledge-Grounded Dialogue Generation with Pre-trained Language Models," in AAAI, 2021.